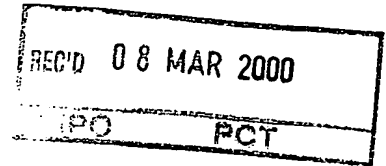


**PRV**PATENT- OCH REGISTRERINGSVERKET  
Patentavdelningen**Intyg  
Certificate**

Härmed intygas att bifogade kopior överensstämmer med de handlingar som ursprungligen ingivits till Patent- och registreringsverket i nedannämnda ansökan.

This is to certify that the annexed is a true copy of the documents as originally filed with the Patent- and Registration Office in connection with the following patent application.

(71) Sökande                      *Lars Gustaf Liljeryd, Solna SE*  
Applicant (s)

(21) Patentansökningsnummer    *9900256-0*  
Patent application number

(86) Ingivningsdatum                      *1999-01-27*  
Date of filing

*Stockholm, 2000-03-02*

*För Patent- och registreringsverket*  
*For the Patent- and Registration Office*

*A. Södervall*  
Anita Södervall

*Avgift*  
Fee

**PRIORITY  
DOCUMENT**  
SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)

## METOD OCH ANORDNING FÖR FÖRBÄTTRING AV EFFEKTIVITET OCH KVALITET HOS LJUDKODARE

### 5 TEKNIKOMRÅDE

Föreliggande uppfinning beskriver metoder och teknik för förbättring av kodningseffektivitet och ljudkvalitet hos ljudkodare och speciellt vid förbättring av ljudkodare som använder SBR-teknik (Spectral Band Replication), en ny teknik som utnyttjar spektralredundans hos ett excitationsspektrum, enligt patentansökningarna Sv Pat 9800268-6 och PCT WO 98/57436. SBR-tekniken ger en markant förbättring av kodningseffektiviteten genom bandbegränsning av signalen vid kodning följt av bandbreddsökning vid avkodning. SBR-tekniken kan utföras antingen som en blind process i avkodaren, SBR-1, eller med hjälp av styrinformationen, i form av spektralenvelopper som skickas från kodaren till avkodaren, SBR-2. Föreliggande uppfinning beskriver ett generellt och effektivt sätt att koda och justera spektralenvelopper. Vidare beskrivs nya tekniker, metoder och förbättringar i de tillämpningar där SBR-teknikens utnyttjas. Uppfinningen kan användas exempelvis inom olika typer av effektiva kodare såsom MPEG 1/2 Layer I/II/III, MPEG 2/4 AAC, Dolby AC-3, NTT TwinVQ, AT&T PAC, CCITT/ITU-R G.722 m.fl. samt till olika former av talkodare och vid konstruktion av nya kodare.

### 20 UPPFINNINGENS BAKGRUND OCH TEKNIKENS STÄNDPUNKT

En linjär PCM signal med CD-kvalitet har 44.1 kHz samplingsfrekvens och 16 bitars upplösning vilket ger ett bitflöde av 1411 kbit/s för en stereosignal. För att reducera bitflödet används psykoakustiska kodare vilka utnyttjar begränsningar i det mänskliga hörselsinnet, sk irrelevans- och redundans-kodning. Härvid kan mer än 90% av informationsmängden i en vanlig CD elimineras utan att lyssnaren upplever någon försämring av ljudkvaliteten. Mycket hög ljudkvalitet kan därigenom erhållas vid c:a 128 kbit/s. Vid lägre kvalitetskrav kan den bästa kodningstekniken i dag återge tal och musikinformation med begränsad ljudkvalitet vid c:a 24 kbps i mono. Vid mycket låga bitflöden erhålls viss distorsion och begränsad bandbredd.

Inom audiokodning delas signalen vanligtvis upp i två komponenter, en spektralenvelopsrepresentation och en tillhörande residuals signal. I bitströmmen utgör residuals signalen huvuddelen av informationsflödet och spektralenveloppen en betydligt mindre del. Vid låga bitflöden minskar detta förhållande varför det är extra viktigt att kunna överföra spektralenveloppen på ett kompakt sätt. Föreliggande patent uppvisar metoder som ger stora förbättringar inom detta område.

Spektralenveloppen är en funktion av två variabler, tid och frekvens. Kodning kan utföras i frekvensled eller i tidsled i det tvådimensionella tid/frekvensplanet. Vanligtvis kodas spektralenveloppen i frekvensled (tex PCM, DPCM, LPC eller vektorkvantisering) vid diskreta  
5 tidpunkter.

Inom "natural-audio" kodning används ofta DPCM vid kodning av spektralenveloppen. DPCM kodningen använder vanligtvis inte någon prediktion som är brukligt i DPCM-sammanhang, utan man kodar skillnaden i amplitud mellan närliggande frekvensvärden (skalfaktorer). Inom applikationer där  
10 låg bitrate används tillsammans med SBR öppnar sig nya möjligheter för kodningen av spektralenveloppen över det frekvensintervall som återskapas. Detta för att spektralenveloppen som skickas kan betraktas mer som styrdata till SBR processen, än bara enveloppen av en signal. Inom t.ex. MPEG Layer 1,2 används 6 bitars linjär PCM kodning av skalfaktorena, detta motsvarar approximativt 1.5 dB steg och ger därför en dynamik av 96dB. I t.ex. Dolby AC3 används  
15 differenskodning av skalfaktorena, dock utan prediktion. Den första skalfaktorn för det lägsta subbandet skickas som ett linjärt PCM värde varefter de följande (ökande i frekvens) beräknas som skillnaden till den angränsande lägre skalfaktorn.

Vid filtrering av residualen vid talkodning används ofta linjär prediktion (LPC). För att nå högsta  
20 kodningseffektivitet kan filterbandbredderna relateras till mänskliga hörselns frekvensupplösning. Vid "natural-audio" kodas spektralenveloppen oftast m h a mångkanaliga filterbankar. Vanligen används filterbankar med konstant bandbredd varvid filterkanalerna grupperas med bark-upplösning. Det finns även filterbankar med konstant relativ bandbredd. Inom talkodningen används "Bark-warped LPC", vilken även kan användas med SBR.

25 Inom de flesta kodningsmetoder extraheras spektralenveloppsinformationen blockvis. Äldre kodningssystem använder konstant blocklängd som ger god temporal upplösning. Detta ger hög kodningskostnad. För att förbättra kodningseffektiviteten byter moderna kodare blocklängd beroende av programmateriallets innehåll. Långa block används för att koda kvasistationära signaler och korta  
30 block används för att få hög temporal upplösning vid transienta förlopp. Uppfinningen beskriver en ny och bättre metod för skalfaktorgruppering i kombination med en indexering som medger kompakt signalering av blockindelningen.

Vid stereokodning används summa/differenskodning (även benämnd M/S-kodning) eller  
35 intensitetsstereokodning. Företrädesvis används summa/differenskodning då denna erbjuder större

- flexibilitet och ljudkvalitet än intensitetsstereo. Vid extremt låg bithastighet måste programaterialet nedsamplas i kodaren vilket ger begränsad audiobandbredd. SBR-tekniken används med fördel för att utöka denna. Då SBR-tekniken normalt används i stereo, d v s separat för både (L och R) eller för både summa (L+R) och differenskanalerna (L-R) kräver kodning av envelopperna en dubblering av spektralenveloppinformationen för SBR. Föreliggande uppfinning beskriver nya metoder för att förbättra stereokodningen.

#### SAMMANFATTNING AV UPPFINNINGEN

- 10 Målet med uppfinningen är att erbjuda nya metoder och anordningar för att öka kodningseffektiviteten eller audiokvaliteten i ett kodningssystem för digital audio. Uppfinningen avser huvudsakligen att förbättra kodningseffektiviteten och kvaliteten av SBR-uppbyggda kodare. Delar av uppfinningen kan även appliceras i andra kodningssammanhang.
- 15 Uppfinningen beskriver en ny metod att gruppera subbandssamplers i tid och frekvens vid generering av skalfaktorer för enveloppskodning. För varje grupp om T samplers i tid och F samplers i frekvens beräknas en skalfaktor, vilken representerar t ex medeleffekten i denna region. Vid kvasistationära signaler används stora värden på T och små värden på F varvid god frekvensupplösning erhålles. I kodaren finns en transientdetektor med vars hjälp transienters positioner fastställs. I omedelbar anslutning till transientens start, påbörjas en ny grupp med ett litet värde på T. Denna följs av ytterligare en eller flera grupper med små värden på T. På detta vis kodas transientens starttidpunkt och utklingingsförlopp med god tidsupplösning. För att kompensera datamängdsökningen härrörande från minskningen av T, väljs större F för dessa grupper. Uppfinningen beskriver hur gruppindelningen kan signaleras med få kontrollbitar. Skalfaktorerna sammanförs till granuler vilka sänds med en konstant uppdateringsfrekvens. Transienternas starttidpunkter beskrivs med ett index som anger positionerna inom dessa granuler. Eftersom sändare och mottagare förfogar över regler som beskriver hur grupperingen skall ske vid ett givet index, räcker det att skicka detta index samt en transientflagga för en entydig avkodning av enveloppen.
- 20 25 30 Uppfinningen utnyttjar vidare en ny typ av kodning av skalfaktorerna. Denna är baserad på en iterativ analys-genom-synteskodning av envelopperna i både tids- och frekvensled. Eftersom en signal inte kan vara helt transient i både tidsdomänen och frekvensdomänen vid en given tidpunkt, kan kodningseffektiviteten vid användandet av alla typer av deltakodning av spektralenvelopper ökas genom att adaptivt växla mellan att koda spektralenveloppen i tids eller frekvensled.

Vid SBR används ofta en kombination av flera transponeringsfaktorer. En metod att utnyttja enveloppsjusterarens analysida för adaptiv nivåjustering mellan de olika deltonsordningarna beskrivs.

Om inga begränsningar för enveloppsjusterarens förstärkningsfaktorer införs, innehåller det replikerade högbandet under vissa omständigheter icke önskvärt smalbandigt brus. Uppfinningen visar ett effektivt sätt att medelst brusgrindar eliminera detta, utan att replikeringen i övrigt påverkas.

För att minimera bitflödet används största möjliga bandbredder på skalfaktorerna. Uppfinningen beskriver hur skalfaktorerna kan interpoleras i frekvens och enveloppsjusteringen operera med dessa smalare bandbredder. Härvid uppnås en större kontroll över det replikerade högbandets spektrum och en effektivare funktion hos brusgrindarna.

En frekvensdomänstransponeringsalgoritm baserad på användandet av filterbankar eller transformer med variabla bandbredder beskrivs.

En förstärkningsfaktorsutjämning för att undvika ringning och vikning i filterbanken som används för enveloppsjustering beskrivs.

Vid SBR kan filterbanksbaserade transponerare användas. Uppfinningen beskriver ett sätt att minimera för- och efterkon hos dessa, m h a översampling i frekvensdomänen.

Ett förbättrat sätt att avkoda stereoinformation med SBR beskrivs.

25

#### KORTFATTAD BESKRIVNING AV RITNINGARNA

Föreliggande uppfinning kommer nu att beskrivas med hjälp av exempel med tillhörande ritningar, vilket ej skall uppfattas som begränsningar av uppfinningens utförande, där:

- Fig. 1 illustrerar ekvidistant tidsindelning av skalfaktorer.
- Fig. 2 illustrerar icke-ekvidistant tidsindelning av skalfaktorer.
- Fig. 3 illustrerar användandet av indexeringsinformation för granulsegmentering.
- Fig. 4 illustrerar enveloppskodning i tids- och frekvensled.
- Fig. 5 är ett flödesschema över T/F-algoritmen.

35

Fig. 6 illustrerar avkodarsystemet.

Fig. 7 illustrerar utjämning av förstärkningsfaktorer.

Fig. 8 illustrerar patchning utan respektive med frekvensdomänsöversampling.

Fig. 9 illustrerar tillvägagångssätt vid nollpadding i en filterbank för transponering.

5 Fig. 10 illustrerar patcher vid fix respektive variabel bandbredd.

Fig. 11 illustrerar tids och frekvensupplösning i en icke-ekvidistant filterbank.

Fig. 12 illustrerar kodaren.

#### DETALJERAD BESKRIVNING AV FÖREDRAGNA UTFÖRINGSFORMER

- 10 Ljudkodning delas vanligen upp i två kategorier, talkodning och "natural-audio" kodning. Talkodning är oftast parametrisk vilket innebär att den är optimerad för tal. "Natural-audio" kodning är icke-parametrisk vilket medför att valfritt ljud kan kodas. Den föreliggande nya metoden för spektralenveloppskodning eller skalfaktorgenerering har framtagits speciellt med tanke på SBR-tillämpningar, men kan framgångsrikt användas för traditionell subbandskodning eller inom andra
- 15 kodningssammanhang, t ex talkodning, där en effektiv representation av spektralenveloppen är önskvärd. Detta inkluderar även applikationer uppbyggda kring SBR-1 konceptet, dvs system där spektralenveloppen inte överförs utan istället estimeras i mottagaren med tex asymptoter, kodböcker baserade på vektorkvantiserade envelopper, system med sk "Statistical Recovery Functions" ["Statistical Recovery of Wideband Speech from Narrowband Speech", Cheng et al, IEEE Transactions
- 20 On Speech and Audio Processing, Vol. 2, No 4, October 94] etc. Även dessa system kan med fördel använda nedan beskrivna interpolering i tid och frekvens, brusgrindning samt icke-uniform sampling av spektralenvelopper.

- I konventionella subbandskodare delas subbandsignalerna upp i skalfaktorer och skalade
- 25 subbandssignaler vid analysen i kodaren. Denna uppdelning kan ske på många olika sätt, och flera kanaler kan använda en gemensam skalfaktor. Gemensamt för alla systemen är att såväl skalade subbandsamplare som tillhörande skalfaktorer överförs och kombineras på något sätt vid syntesen i dekodern. Vid SBR är detta inte fallet, endast signalens spektrala grovstruktur behöver överföras, vilket hos vissa kodare skulle motsvara överföring endast av skalfaktorerna. Detta ställer nya, hårdare
- 30 krav på hur skalfaktorerna genereras eftersom inga skalade subbandsamplare, uppbärande information om småskaliga tidsförlopp, finns att tillgå. Problematiken belyses med ett exempel:

- I fig. 1 visas tid-/frekvensrepresentationen av en musiksingal där en organist spelar uthållna ackord och en trumslagare ackompanjerar på en hi-hat. I lågbandet har orgeln hög energi och hi-haten låg,
- 35 medan förhållandet i högbandet är det omvända. De skalfaktorer som genereras inom de tidsintervall

där hi-haten är närvarande, domineras av dennas höga intermittenta energi. Vid SBR-processen i dekodern estimeras enveloppen hos den transponerade signalen synkront med den kodade originalenveloppen, med samma tids- och frekvensupplösning, fig 1. Förstärkningsfaktorena hos filterbanken för enveloppsjustering erhålles ur kvoten mellan de båda estimaten. Härvid uppstår för denna typ av signal ett problem: Den transponerade signalen har samma energiförhållande mellan hi-hat och orgel som lågbandet. De förstärkningar som krävs för att styra den transponerade hi-haten till rätt nivå orsakar därför att den transponerade orgeln förstärks i förhållande till dess ursprungliga högbandsnivå under den fulla durationen hos den skalfaktor som innehåller hi-hat energi. Denna momentant alltför starka orgel upplevs som för- och efterkon till hi-haten. Denna typ av distorsion kallas skalfaktorinducerat eko. Detta fenomen kan elimineras genom att skalfaktorena uppdateras så ofta i tiden att tillräckligt kort tid mellan uppdatering och transient garanteras för godtyckligt lokaliserade transienter. Ett sådant förfarande skulle förstås avsevärt öka datamängden och är således opraktiskt.

Därför visas ett nytt system som medger en flexibel skalfaktorgenerering och löser ovanstående problem utan att överflödiga data genereras, under användande av ett minimum av kontrollsignaler. Principlösningen ligger i att bibehålla en låg samplingsfrekvens av enveloppen under tonala passager, vilka utgör merparten av ett normalt programmaterial, och att man ha en transientdetektor i kodaren fastställa tidpunkterna för transienterna och uppdatera skalfaktorena i omedelbar närhet till dessas "framkant" eller startflank, se fig. 2. Detta eliminerar skalfaktorinducerade förekon. För att väl reproducera transientens utklingningsförlopp, ökas uppdateringsfrekvensen temporärt under ett tidsintervall efter transientens start. Detta eliminerar skalfaktorinducerade efterkon. Tidssegmenteringen under utklingningen är inte lika kritisk som när det gäller att finna transientens start, eftersom hörselsinnets temporala postmaskering är större än dess premaskering. I syfte att kompensera den ökade datamängd detta förfarande medför, reduceras samplingstätheten i frekvens under transienten.

Notera att denna varierande sampling i tid och frekvens utförs genom att subbandsamplerna från en och samma filterbank grupperas på olika vis. Variabel skalfaktorupplösning i tid och frekvens förekommer även hos vissa konventionella subbandskodare. Skillnaden ligger däri att dessa dels växlar (mellan olika grupperingar) enligt andra kriterier, dels i allmänhet också byter storlek på filterbanken. En sådan ändring av storleken kan inte ske omedelbart utan måste ske med hjälp av övergångsfönster, varför uppdateringstidpunkterna inte kan väljas lika fritt som om filterbanken förblir oförändrad. Vidare önskas vid SBR en relativt detaljerad kodning av enveloppen hos de tonala signalerna ända fram till det ögonblick då transienten inträffar. Även detta krav tillgodoses genom att

filterbankstorleken hålls konstant eftersom vid varje tidpunkt denna banks goda upplösning i frekvens finns att tillgå.

För att avkodaren skall veta hur skalfaktorena har grupperats, krävs någon form av signalering. Denna måste vara flexibel nog att täcka alla kombinationer av intresse, utan att generera en alltför stor mängd styrdata. Typiska kodare arbetar på en blockbasis, där varje block motsvarar ett fixerat tidsintervall. Om en icke-uniform sampling enligt fig. 2 skall kunna realiseras i ett sådant system, måste alltså även problemet med segment som spänner över blockgränserna hanteras. En ny metod för indexering av transienter samt implicit signalering av de olika grupperingsfallen beskrivs nu.

Antag att ett block uppdelas i  $2^N$  underblock i tid. Positionen för ett underblock kan då anges med  $N$  bitar. En bit åtgår för att signalera förekomsten av en transient inom ett block och  $N$  bitar för att ange positionen i underblock för denna transient. Dessa värden används tillsammans med värden från tidigare block för uppslagning i en tabell som anger hur aktuellt block skall indelas. Om både kodare och avkodare förfogar över samma tabell över skalfaktorens frekvensupplösning samt en uppsättning regler för hur segmenteringen enligt ovan sker vid givna transientpositioner, räcker dessa två signaler för en entydig avkodning. Vid tillämpningar där man kan tillåta en viss fördröjning kan man låta transientdetektorn arbeta i förväg innan man med tabellen bestämmer indelningen av aktuellt block. På detta sätt kan man också göra tidsindelningen mindre beroende av de naturliga gränserna mellan konsekutiva block. Betrakta blocket i fig. 3a som är uppdelat i åtta underblock. Fyra underblock från ett tidigare samt nästkommande block beaktas också. Transientdetektorn, som arbetar med samma upplösning men ett halvt block i förväg, fig. 3b, har tidigare (vid tiden  $n-1$ ) påvisat förekomsten av en transient i underblock 2, vilket motsvarar underblock 6 i block  $n-1$ . I aktuellt block har en transient detekterats i underblock 3 (motsvarande underblock 7 i block  $n$ ). Med dessa positioner som ingångsvärden till tabelluppslagningen fås som exempel att block  $n$  skall indelas som visas i fig. 3c. Block  $n-1$  har tidigare bearbetat det 6:e transienta underblocket, men lämnat det 7:e underblocket, vilket tas om hand i block  $n$ . Förutom den uppenbara fördelen med att tillåta segment som spänner över blockgränserna, kan detta system underlätta uppfyllandet av vissa kriterier, såsom tex kravet att bitflödet för enveloppsdatat skall vara konstant.

Inom Fourieranalysen är följande samband välkända:

$$\mathcal{F}[\delta(t)] = 1$$

$$\mathcal{F}[1] = 2\pi\delta(\omega)$$



Vilket innebär att en puls i tidsdomänen motsvaras av ett flackt spektrum i frekvensdomänen, och en "puls" i frekvensdomänen motsvaras av en stationär signal i tidsdomänen med en frekvens motsvarande pulsens läge i spektrat. Detta kan tolkas som att en signal aldrig kan vara transient i två domäner samtidigt. Ovanstående samband kan med fördel utnyttjas inom kodning av

5 spektralenvolupper eftersom en tonande stationär signal kan ha ett mycket spretigt frekvensspektrum icke lämpat för deltakodning i frekvensled, samtidigt som spektralenvoluppen inte ändras märkbart i tid. Detta illustreras i fig. 4.

I en implementation av den föreslagna uppfinningen används ADPCM-kodning för att koda

10 skalfaktorerna. Denna kodning använder en fast prediktor i form av två uppslagstabeller, en indextabell och en stegtabell. Detta gör att algoritmen kan implementeras mycket beräkningseffektivt, vilket gör en iterativ analys-genom-syntes kodning möjlig. I motsats till den DPCM kodning som används i Dolby AC3 kodas här inte skillnaden i nivå mellan skalfaktorerna utan en indexskillnad i den tabell där amplitudvärdena finns tabellerade, stegtabellen.

15 Följande exempel kommer att användas för att förklara uppfinningens natur utan att för den skull begränsa den. Vidare kommer i den följande utläggningen spektralenvoluppen representeras av en vektor med skalfaktorer vid tidpunkten  $n_0$

$$Y(k, n_0) = [a_1, a_2, a_3, \dots, a_k, \dots, a_N],$$

20 där  $a_1 \dots a_N$  är enveloppens amplitudvärden vid olika frekvenser. Det vanliga angreppssättet är att koda skillnaden mellan närliggande värden vilket ger en vektor

$$D(k, n_0) = [a_2 - a_1, a_3 - a_2, \dots, a_N - a_{(N-1)}].$$

För att kunna avkoda denna krävs även att startvärdet  $a_1$  skickas med. Som nämndes ovan kan detta visa sig mycket ineffektivt om spektrat innehåller endast ett par stationära toner eftersom det då får en

25 mycket transient natur. Detta kan göra deltakodningen i frekvensled dyrare än vanlig PCM kodning. För att komma tillrätta med det problemet föreslås här en switchad kodning, som kodar både i tids- och frekvensled och beräknar kodningsfelet i båda riktningar för ett givet antal bitar, och avgör på så sätt vilken riktning som är mest fördelaktig att koda i.

30 Den föreslagna algoritmen, härafter kallad T/F-kodning, är en iterativ analys-genom-syntes-algoritm. I en DPCM implementering beräknas två delta vektorer,  $D_f$  och  $D_t$ :

$$D_f(k, n_0) = [a_2 - a_1, a_3 - a_2, \dots, a_N - a_{(N-1)}]$$

$$D_t(k, n_0) = [a_1(n_0) - a_1(n_0 - 1), a_2(n_0) - a_2(n_0 - 1), \dots, a_N(n_0) - a_N(n_0 - 1)]$$

Dessa kodas med lika många bitar per värde, varefter de avkodas. Felvektorer beräknas för de två kodningsvarianterna

$$E_f(k, n_0) = [e_1, e_2, \dots, e_N] = Y(k, n_0) - Y_{f_{est}}(k, n_0)$$

$$E_t(k, n_0) = [e_1, e_2, \dots, e_N] = Y(k, n_0) - Y_{t_{est}}(k, n_0)$$

- 5 Ett skalärt värde på kodningsfelet, eller spektraldistorsionen, kan beräknas på flera olika sätt: antingen används maxavvikelsen inom något frekvensband eller ett medelfel över alla kodade skalfaktorer.

$$dist(n_0) = \max_{k=1}^N |E(k, n_0)|$$

$$dist(n_0) = \frac{1}{N} \sum_{k=1}^N |E(k, n_0)|$$

- Den kodade-avkodade vektor med det minsta felet representerar den mest fördelaktiga riktningen att koda spektralenvelopen i vid den givna tidpunkten. Om felet understiger det av kodaren förbestämda maxfelet skickas den kodade enveloppen, annars ökas bitantalet med vilken envelopperna kodas och proceduren upprepas tills maxfelet understigs eller tillåtet max antal bitar uppnåts. Proceduren beskrivs i flödesschemat i fig. 5.
- 10

- I den nuvarande implementeringen används ADPCM-kodning, det skall dock beaktas att alla sorters differenskodning med fördel kan användas i den föreslagna T/F-kodningsalgoritmen. Vid kodning med ADPCM kan man låta antalet bitar som används för att koda skalfaktorerna variera på skalfaktor- eller enveloppsbasis. Båda varianterna har sina för och nackdelar som beskrivs nedan. Används konstant antal bitar kan tabellerna som används för ADPCM kodningen utformas mer optimalt. Detta innebär att ett indexdelta för fyra bitars ADPCM inte är ekvivalent med samma indexdelta för
- 20 ADPCM med tre bitar, vilket vidare innebär att antalet bitar som har använts för att koda skalfaktorerna måste skickas som sidoinformation till enveloppsdatat. Nackdelen är att om enveloppsvärdena alla kräver få bitar utom ett som kräver många kommer alla värden att kodas med det högre antalet bitar, vilket kan bli ineffektivt. Alternativt kan kodningen göras med variabelt antal bitar på skalfaktorbasis. Sådan kodning (utan T/F switchning) är den vanliga inom deltakodning i
- 25 audiokodare. Fördelen är att varje sampel kodas endast med så många bitar som krävs för att representera värdet. Används ADPCM-kodning måste tabellerna göras kompatibla så t.ex. indexdelta = 2 är ekvivalent oavsett hur många bitar som använts att koda värdet med. Det ska dock beaktas att till detta måste en synkroniserande kodning adderas såsom Huffman för att mottagaren ska kunna tolka datat korrekt. Vidare är inte alltid Huffman att föredra eftersom en sådan typ av
- 30 redundanskodning förutsätter en viss statistisk fördelning av signalen som inte nödvändigtvis existerar. Startvärdena som skickas när deltakodning sker i frekvensled behöver naturligtvis inte skickas då delta-kodningen sker i tidsled eftersom de redan är tillgängliga i form av den föregående enveloppen.

ADPCM-kodningen såsom använd i den föreslagna algoritmen gör det möjligt att koda med olika amplitudupplösning för olika frekvenser, eftersom algoritmen opererar på kvantifierade tabelldata, och endast kodar tabellsteg, istället för det verkliga datat. Detta innebär att precis som bandbredden kan ökas för respektive skalfaktorband i förhållande till barkbandbredden kan även amplitudupplösningen justeras efter örats frekvensupplösning.

Den föreslagna algoritmen kräver att ytterligare sidoinformation skickas tillsammans med den kodade enveloppen. Detta extradata består av en tid-frekvensflagga som signalerar i vilken riktning spektralenveloppen har kodats, vidare måste antal bitar eventuellt, som ovan beskrivits, även skickas med. Detta extra data är dock obetydligt i förhållande till den kodningvinst som algoritmen ger.

Vid kodning av spektralenvelopper för SBR är omständigheterna något annorlunda än för vanlig enveloppskodning. Den mottagna spektralreplikerade signalen har en formantstruktur och envelopp, skapad av transponeraren. Den mottagna enveloppen skall användas för att justera den befintliga. Det är då möjligt att utnyttja redundans mellan lågbandet (hämtområdet) och högbandet, dvs istället för att deltakoda närliggande frekvenssampel kodas frekvenssamplerna på oktavbasis.

En ideal SBR process skapar med många transponeringsfaktorer ett stort antal harmoniska komponenter, och ger samma harmoniska densitet som originalet. Här beskrivs en metod att välja lämpliga förstärkningar till de olika deltonsordningarna.

Låt insignalen vara en harmonisk serie

$$x(t) = \sum_{i=0}^{N-1} a_i \cos(2\pi f_i t)$$

Vid en transponering en faktor 2 erhålles

$$y(t) = \sum_{i=0}^{N-1} a_i \cos(2 \times 2\pi f_i t)$$

Som synes saknas varannan delton i ett på detta vis skapat högband. För att "fylla på" med deltoner, summeras transponeringar med andra faktorer, t ex  $M = 3, 5$ , till andratonen enligt ovan. För att dra maximal nytta av multipla deltonsordningar är det viktigt att de matchas väl i nivå så att ingen dominerar över de andra (inom överlappande frekvensområden). En svårighet härvidlag är att

signalnivån kan variera starkt mellan de olika transponerarnas käll-/hämtområden. Dessa skillnader varierar dessutom mellan olika typer av programmaterial, varför det är svårt att använda fasta förstärkningar för transponerarna. En metod för nivåanpassning som tar hänsyn till insignalens spektralfördelning i lågbandet framgår av fig. 6. Transponerarnas utsignaler adderas och skickas till filterbanken för enveloppsjustering. Till denna summa läggs även lågbandssignalen varvid filterbankens analysida kan användas för spektralanalys även av lågbandet. Tex kan totala effekten inom respektive hämtområde bestämmas och transponerarförstärkningarna justeras motsvarande. Det är också möjligt att estimeras spektrallutningar och kompensera motsvarande före filterbanken med tex enkla filter av shelving-typ. Notera att detta förfarande inte påverkar filterbankens grundfunktion som equalizer och att lågbanddelen av analysignalen inte skickas till filterbankens syntessida.

Enligt resonemanget ovan innehåller det syntetiskt genererade högbandet ibland hål i spektrat. Enveloppsjusteraren söker att justera den transponerade signalen så att en likadan spektralenvelopp som originalets erhålles. Antag att originalsignalen har hög energi inom ett visst skalfaktorband och att den transponerade signalen uppvisar ett sådant spektralt hål i detta frekvensområde. Om förstärkningarna tillåts att variera fritt, kommer en mycket hög förstärkning att erhållas inom detta band och brus eller signalrester att lyftas upp till originalsignalens nivå. Detta fenomen kallas icke önskvärd brussubstitution. Låt  $P_1 = [p_{11}, \dots, p_{1N}]$  vara originalsignalens skalfaktorer vid en viss tidpunkt och  $P_2 = [p_{21}, \dots, p_{2N}]$  den transponerade signalens dito, där varje element i vektorerna representerar en subbandsenergi normerad mot p tid och frekvens. Erforderliga förstärkningsfaktorer i filterbanken för enveloppsjustering ges då av  $G = [g_1, \dots, g_N] = [\sqrt{p_{11}/p_{21}}, \dots, \sqrt{p_{1N}/p_{2N}}]$ . Genom att studera  $G$ , kan lätt subband med brussubstitution identifieras; dessa kännetecknas av att motsvarande förstärkningar är mycket högre än övriga subbands. Brussubstitution kan enkelt undvikas genom att förstärkningarna limiteras, dvs tillåts att variera fritt endast upp till ett visst maxvärde,  $g_{\max}$ . Förstärkningarna med "brusgrindar" ges alltså av  $G_{\text{gate}} = [\min(g_1, g_{\max}), \dots, \min(g_N, g_{\max})]$ . Detta uttryck illustrerar emellertid endast grundprincipen för brusgrindarna. Eftersom originalenveloppen och den transponerade signalens envelopp kan skilja mycket i nivå och lutning, kan inte fasta maxvärden  $g_{\max}$  användas. Istället beräknas en medelförstärkning och förstärkningarna tillåts att överskrida denna med en viss faktor. För att ta hänsyn till storskaliga (bredbandiga) nivåvariationer hos envelopperna, kan vektorerna  $P_1$  och  $P_2$  delas in i segment, var och ett innehållande ett större antal skalfaktorer, som behandlas individuellt enligt ovan. På detta sätt kan en mycket effektiv grindfunktion erhållas utan att inskränka funktionen av nivåjusteringen av de subband som innehåller nyttsignaler.

Som har beskrivits ovan, grupperas kodarens analyskanaler i frekvens enligt olika tabeller vid beräkningen av skalfaktorer. Varje sådan skalfaktor utgör ett estimat av spektraltätheten inom det

- frekvensband som spänns upp av de grupperade kanalerna. För att erhålla så lågt bitflöde som möjligt vill man naturligtvis minimera antalet skalfaktorer, varför så stora kanalgrupper som möjligt eftersträvas. En vanlig lösning inom audiokodning är att anpassa skalfaktorbandbredderna efter en bark-skala, varvid hörselsinnet i huvudsak logaritmiska frekvensupplösning utnyttjas. I en SBR-
- 5 mottagares filterbank för enveloppsjustering kan, i likhet med traditionella kodare, en likadan gruppering av kanalerna ske, så att frekvensjustering sker med samma filterbandbredder som används vid skalfaktorgenereringen. Emellertid kan kanalerna i mottagarens filterbank tillåtas operera individuellt genom att en skalfaktor per filterbankskanal interpoleras fram ur de mottagna skalfaktorerna. Det enklaste sättet att göra detta, är att tilldela alla kanaler som ingick vid beräkningen
- 10 av en skalfaktor just detta värde. Den transponerade signalen analyseras varvid en skalfaktor per kanal beräknas. Dessa faktorer samt de interpolerade skalfaktorerna från originalsignalen används för att beräkna kanalförstärkningar enligt ovan. Denna frekvensdomänsinterpolering har två fördelar. Den transponerade signalen har (som förklarats tidigare) i allmänhet ett glesare spektrum än originalet. En spektral utjämning fordras då och denna blir effektivare än om de förhållandevis stora
- 15 skalfaktorbandbredderna används i equalizern. M a o, de syntetiskt genererade övertonerna isoleras bättre och kan styras mer individuellt mot önskad nivå inom ett visst frekvensområde. Vidare blir funktionen hos brusgrinden bättre, eftersom brusgolvet kan isoleras och dämpas med större precision i frekvens.
- 20 Skalfaktorer eller spektralenveloppsvärden estimeras i kodaren genom att beräkna toppvärdet i eller medelvärde över en matris i tid/frekvens-planet, där de ingående värdena vanligtvis är kvadrerade komplexvärda subbandssampel. Matrisernas kolumnantal beror på aktuell tidsupplösning och radantalet beror på aktuell frekvensupplösning. Båda dessa värden ändras adaptivt beroende på signalens statistik. Matriser med få rader och många kolumner (hög frekvensupplösning och låg
- 25 tidsupplösning) används vid tonala förlopp, och motsatsen gäller vid transienta passager. För att höja kodningseffektiviteten är oftast frekvensupplösningen på något sätt anpassad till den mänskliga hörseln. Frekvensupplösningar med nära konstant relativ bandbredd, såsom Mel- eller bark-relaterade upplösningar används därför med fördel. På samma sätt kan man också låta tidsupplösningen variera med frekvensen, med lägre tidsupplösning för lägre frekvenser. I båda dessa fall kan den
- 30 underliggande filterbanken vara baserad på Wavelet eller Famlet teknik, där filterbankar med variabel tid/frekvens-upplösning kan erhållas.

- I avkodaren beräknas på samma sätt som i kodaren spektralenveloppsvärden över matriserna i tid/frekvens-planet. De ingående värdena är nu subbandssampel erhållna från det SBR-replikerade
- 35 frekvensområdet. Avkodaren beräknar förhållandet mellan de i kodaren och avkodaren estimerade

skalfaktorerna och applicerar dessa förstärkningsfaktorer på subbandssamplen för att på så sätt justera spektralenvelopen för det SBR-replikerade frekvensområdet. Stora skillnader mellan närliggande förstärkningsfaktorer, i både tid och frekvens, kan resultera i ringning och vikning i den justerande filterbanken. Det är då fördelaktigt att göra dessa språng mindre skarpa genom utjämning av

5 förstärkningsfaktorer. Fig. 7a visar förstärkningsfaktorer som skall multipliceras med motsvarande subbandsamplar. I figuren ses att två högupplösande block följs av två lågupplösande och ett högupplösande, och även att bandbredden på skalfaktorbanden ökar exponentiellt med ökande frekvens. De skarpa sprången i fig. 7a är eliminerade i fig 7b genom filterering av

10 förstärkningsfaktorer i både tid- och frekvensled med ett lågpasfilter. Det är viktigt att skarpa flanker i tid härrörande från de korta lågupplösande blocken inte undertrycks alltför mycket, då detta skulle försämra det SBR-replikerade områdets transientenergi. Detsamma gäller för filtreringen i frekvensled över de högupplösande långa blocken, då en alltför tillplattad envelopp leder till en bristfällig formantstruktur. I fig 7b är filtreringen medvetet överdriven.

15 I patentansökan SV Pat. 9800268-6 och PCT Pat. WO 98/57436 beskrivs hur transponering kan utföras m h a filterbanker (t ex STFT eller MDCT) genom att ansluta analyskanaler till synteskanaler på ett nytt sätt. En sådan mappning av kanaler gavs där namnet "patch". Som framgår av figuren, har patcherna för transponering det gemensamt att synteskanaler lämnas utan insignal. Transponeringen blir exakt endast för en diskret frekvens inom varje analyskanals passband. I patentansökan # beskrivs

20 hur fasen hos de komplexvärda subbandssignalerna som ingår i patchen kan modifieras så att transponeringen blir exakt för insignaler med godtycklig frekvens (inom det frekvensområde transponeraren har dimensionerats för). En transponering enligt detta kan ses som en process i två steg där patchen står för en grovskalig, approximativ transponering och blocken P kan ses som lokala transponerare som opererar på de lågfrekventa bandpasssignalerna. Denna lokala transponering

25 fungerar endast då varje analyskanal innehåller högst en diskret frekvens. Om insignalen är en dirac-puls ges analysignalerna av impulsresponsen hos filterbankens prototypfilter. Denna typ av signal kan inte transponeras (bandbreddsutökas) lokalt av P-blocken och utsignalen får för- och efterekon, eftersom impulsresponser fattas vid syntesen. I patentansökan SV Pat. 9800268-6 och PCT Pat. WO 98/57436 föreslogs ett antal metoder att komma runt detta problem, men här visas en ny metod:

30

I fig 8a visas symboliskt frekvensdomänsrepresentationen av analys- och syntessignalen för en patch då insignalen är en dirac-puls och transformstorleken är  $N$ . Antag att syntesen normalt opererar på segment om  $L$  samplers. Istället bildas ett nytt segment med längden  $2L$  där samma  $L$  samplers som tidigare ingår och resten utgörs av nollor. Detta förfarande är vanligt vid spektralestimering och

35 benämns nollpadding. Segmentet med längden  $2L$  transformeras med en  $2N$ -transform vilket

resulterar i en frekvensdomänsrepresentation som är två gånger översamplad i frekvens. Det är viktigt att notera att en dylik frekvensdomänsöversampling inte ökar frekvensupplösningen hos analysen, utan endast motsvarar en tätare evaluering av transformen. I fig 8b visas analys- och syntessignalerna vid nollpadding. Vid analysignalen föreligger som synes överlappning vilken "dras isär" av patchen, varför det vid syntes finns en komplett uppsättning impulssvar, trots att varannan synteskanal alltså saknar insignal. Eftersom hela utfrekvensintervallet spänns upp av dessa impulssvar, förekommer inte längre några för- eller efterekon.

Metoden att översampla i frekvens kan generaliseras till andra transponeringsfaktorer än två. Vid en transponering  $M$  fordras även en översamlingsfaktor  $M$ . Vid fallet STFT kan vid översampling  $M$  gånger alla transponeringar till och med  $M$  erhållas utan ekon, d v s fyra gångers översampling medger samtidig generering av två, tre och fyra gångers transponering om så önskas. Det skall påpekas att överlappningen i frekvens hos analyskanalerna inte inverkar på funktionen hos P-blocken. En och samma frekvenskomponent hos insignalen finns på detta sätt representerad i två eller flera analyskanaler, men P-blocken och patchen säkerställer att dessa olika instanser hamnar på samma utfrekvens.

I en filterbanksimplementation multipliceras insignalen på analysidan på vanligt sätt med ett fönster (prototypfilter) som flyttas ett visst antal sampel mellan varje multiplikation, fig. 9. Vid kritisk sampling flyttas fönstret typiskt  $L/M$  sampel, där  $L$  är antalet kanaler eller frekvenskoefficienter och  $M$  är transponeringsfaktorn. Ett fönstrat block om  $2L$  sampel paddas därefter med  $2(M-1)L$  nollsampel där hälften av dessa läggs på före respektive efter de fönstrade samplen. Dessa  $2ML$  sampel komplexmoduleras och genererar sålunda  $ML$  frekvenskoefficienter. Denna filterbank är alltså översamplad i frekvens med en faktor  $M$ . På syntessidan demoduleras de  $ML$  frekvenskoefficienterna och de första respektive sista  $(M-1)L$  samplen kasseras.

Frekvensdomänstransponeringsalgoritmerna som presenterades som exempel i SV Pat. 9800268-6 och PCT Pat. WO 98/57436, byggde på användandet av filterbankar eller transformer med konstanta bandbredder. För att tillåta avvägningar mellan olika tid- och frekvensupplösningar kan filterbankar eller transformer med variabla bandbredder istället användas. För förbättrade transientegenskaper föreslås nu en transponerare baserad på "Wavelet packets". Wavelet packets kan betraktas som träduppdelade ortonormala filterbankar med godtycklig struktur. I transponeraren används komplexvärda Wavelets och "scaling functions". Med den godtyckliga uppdelningen kan filterbankar med oktavbandsindelning eller andra filterbankar med variabel bandbredd konstrueras. I fig. 10 visas en transponeringspatch för en filterbank med konstant bandbredd jämfört med patchen för en

filterbank med variabel (oktavbands) upplösning. De filterband som har större bandbredd samplas oftare än de med mindre bredd, varför de förstnämnda har bättre tidsupplösning. Träduppdelningen kan antingen vara statisk eller bestämmas adaptivt under gång, baserat på en sk "best-base" algoritm. En typisk struktur för tid/frekvens-uppdelningen i en statisk Wavelet-packet-transponerare visas i fig. 11.

- 5 De streckade blocken visar exciterade subband för ett kortvarigt transient stimuli. Den minskande frekvensupplösningen för högre frekvenser bidrar till transponerarens goda tidsegenskaper. För lägre frekvenser garanterar en hög frekvensupplösning god kvalitet på transponeringen av tonala signaler.

- Det går att kombinera SBR-tekniken med parametriska analys- och syntesmetoder såsom sinusodialkodning. I dessa fall används pitch-detektering i dekodern för att styra tongeneratorer. 10 Noggrann estimering av pitchen är viktig för att undvika dissonans. Detta förfarande inom SBR-tekniken kallas harmonisk substitution och är ett nytt sätt att använda sinus-analys/syntes tekniken. I de fall fas-komponenten varierar slumpmässigt hos syntesgeneratorm kan denna teknik även användas för att ersätta smalbandiga bruskomponenter, tex hos transienter.

- 15 Som påpekats ovan bör spektral enveloppen som skickas till det med SBR återskapade frekvensområdet, betraktas som ett sorts styrdata snarare än en vanlig spektral envelopp. Detta inbjuder till att i kodaren göra analys-genom-syntes av spektralenvelopen ur ett SBR perspektiv. Enveloppsstyrdatan kan då optimeras så att det SBR-replikerade frekvensområdet efterliknar originalet 20 då styrdatan appliceras. Detta gör att istället för att koda hela högbandsenveloppen så exakt som möjligt kan justeringssignalen kodas mer optimalt eftersom man då utnyttjar eventuell redundans mellan högbandets och lågbandets envelopp.

- Det har nyligen framförts önskemål inom digital rundradio om att kunna sända i stereo på digitala 25 AM-bandet. Bitflödet är så lågt i dessa tillämpningar att audiosignalen måste nedsamplas till låg audiobandbredd. Vid stereokodning används med fördel M/S-formatet dvs summa/differans kodning. Därvid kodas summasignalen respektive differenssignalen separat. Differenssignalen, vilken innehåller mindre information än summasignalen, kodas med färre bitar än summasignalen. Differenssignalen kan även kodas med lägre kvalitet då den främst innehåller "ambiens" och "stereopanorama". Samma 30 förfarande kan tillgripas vid intensitetsstereokodning. Här kodas subbandsamplerna i mono medan skalfaktorerna kodas i stereo. Av bitkostnadsskäl kan det vara nödvändigt att endast överföra SBR-enveloppen i mono. Detta innebär att det replikerade högbandet är i mono, vilket är godtagbart då det mänskliga riktningshörandet är begränsat vid högre frekvenser, fig. 12.



För att förbättra kodningseffektiviteten och ljudkvaliteten hos differenskodaren kan differenssignalen högpPASSfiltreras. HögpPASSfiltreringen medger att fler bitar kan allokeras av kodaren till mellanregistret. Vid inspelning av program panoreras bas-information tex basgitarr och bastrumma oftast till mitten av stereopanoramats. Vidare har hörselsinnet låg riktningsdiskriminering vid låga frekvenser varför det påverkar programmaterialet minimalt då differanssignalen högpPASSfiltreras. Basinformationen överförs i stället i den summakodade kanalen, fig. 13.

En förbättring av stereo-intrycket erhålles då SBR signalen leds genom en ambiensenhet liknande den som beskrivs i SV Pat. 9800268-6 och PCT Pat. WO 98/57436. En ny ambiensenhet för att förbättra upplevd ljudkvalitet beskrivs här. I denna krävs endast en fördröjningsledning, dessutom blir enheten monokompatibel vilket är önskvärt i rundradiosammanhang. Den högfrekventa SBR-signalen erhåller bredd och ambiens. Detta förfarande fungerar lika väl vid bandbredds begränsad intensitetsstereo, fig 14.

Ett intressant alternativ till ovanstående är att använda två SBR-transponerare för att öka bandbredden vid M/S eller intensitetsstereo. Härvid tappas respektive kanal av på (L och R) information och leds till respektive SBR-transponerare. Därefter adderas SBR-signalerna till respektive kanal (L och R) för att erhålla ökad skillnad mellan L och R varvid SBR-signalen justeras av summa-enveloppen (L+R). Även i detta fall kan ambiensenheten användas. För minskad beräkningskomplexitet kan en SBR transponerarenhet användas på mono signalen, varefter högbandet justeras individuellt för de två kanalerna, så att skillnaden i högbandets spektralenvolopp mellan de två kanalerna, är proportionell mot skillnaden i lågbandets spektralenvolopp.

#### Apparatbeskrivning

I fig. 15 visas kodningssidan av den föreslagna uppfinningen. En analys görs på den bredbandiga insignalen varvid spektralenvolppsdata vid ekvidistanta tidpunkter genereras. En transientdetektor skickar sedan styrdata till skalfaktorgenereringsmodulen som med hjälp av återkoppling från en inbyggd SBR-avkodare producerar en icke-uniformt samplad skalfaktorrepresentation av spektralenvolpper. Dessa kodas med T/F-kodning och multiplexas in i bitströmmen tillsammans med bitströmmen från den psykoakustiska kodaren och diverse styrdata.

I fig 6 visas avkodningssidan av den föreslagna uppfinningen. Det mottagna enveloppsdatat avkodas med hjälp av styrdata. Den av ljudkodaren avkodade Lp-signalen uppsamlas och transponeras. Transponerarna använder i vissa fall styrdata från kodaren (TD-transponeraren SV Pat. 9800268-6 och PCT Pat. WO 98/57436) eller översampling i frekvens. Spektralenvolppsdatat interpoleras och

brusgrindning utförs mha det SBR-replikerade högbandet. Det skall poängteras att spektralenveloppsdatat inte nödvändigtvis skickats från kodaren utan kan ha estimerats vid avkodaren. Det replikerade högbandet justeras och en bredbandssignal kommer ut.

## 5 PATENTKRAV

1. Förfarande för bitflödesreduktion vid överföring av spektralenveloppinformation från en sändare till en mottagare, **kännetecknat av** gruppering av subbandssampler i grupper om T sampler i tid och F sampler i frekvens vid generering av skalfaktorer för enveloppskodning. I kodaren finns en transientdetektor med vars hjälp transienters positioner fastställs. I omedelbar anslutning till transientens start, påbörjas en ny grupp med ett litet värde på T. Denna följs av ytterligare en eller flera grupper med små värden på T. Skalfaktorerna sammanförs till granuler vilka sänds med en konstant uppdateringsfrekvens. Transienternas starttidpunkter beskrivs med ett index som anger positionerna inom dessa granuler. Sändare och mottagare förfogar över regler som beskriver hur grupperingen skall ske vid ett givet index, räcker det att skicka detta index samt en transientflagga för en entydig avkodning av enveloppen.

2. Förfarande enligt krav 1, **kännetecknat av** att spektralenveloppinformationen kodas växelvis i tidsled och frekvensled.

3. Förfarande enligt krav 1, **kännetecknat av** en metod att utnyttja enveloppsjusterarens analysida för adaptiv nivåjustering mellan de olika deltonsordningarna beskrivs.

4 Förfarande enligt krav 1, **kännetecknat av** att brusgrindar eliminerar icke önskvärt smalbandigt brus genom att begränsningar för enveloppsjusterarens förstärkningsfaktorer införs

5. Förfarande enligt krav 1, **kännetecknat av** att skalfaktorerna interpoleras i frekvens och enveloppsjusteringen opererar med dessa smalare bandbredder.

6. Förfarande enligt krav 1, **kännetecknat av** en frekvensdomänstransponeringsalgoritm baserad på användandet av filterbankar eller transformer med variabla bandbredder.

7. Förfarande enligt krav 1, **kännetecknat av** en förstärkningsfaktorsutjämning för att undvika ringning och vikning i filterbanken som används för enveloppsjustering.

8. Förfarande enligt krav 1, **kännetecknat** av ett sätt att minimera för- och efterkon hos frekvensdomänstransponerare, m h a översampling i frekvensdomänen.

9. Förfarande enligt krav 1, **kännetecknat** av att SBR i enbart summasignalen (L+R) vid  
5 sum/diff- alt intensitets-stereo används.

10. Förfarande enligt krav 1, **kännetecknat** av att SBR pseudo-stereo enhancement i enbart summasignalen (L+R) vid sum/diff- alt intensitets-stereo används.

10 11. Förfarande enligt krav 1, **kännetecknat** av att SBR aktiv panorama-stereo enhancement i enbart summasignalen (L+R) vid sum/diff- alt intensitets-stereo används.

## Sammandrag

5

10

METOD OCH ANORDNING FÖR FÖRBÄTTRING AV EFFEKTIVITET OCH LJUDKVALITET  
HOS LJUDKODARE

15

20

25

30

35

Föreliggande uppfinning beskriver metoder och teknik för förbättring av kodningseffektivitet och ljudkvalitet hos ljudkodare, speciellt sådana som använder SBR-teknik (Spectral Band Replication, en ny teknik som utnyttjar spektralredundans hos ett excitationsspektrum). Ljudkodare överför spektralenveloppsinformation, från kodaren till avkodaren. Föreliggande uppfinning beskriver ett nytt förfarande att koda och justera spektralenveloppen på ett effektivt sätt vilket även är lämpat för tillämpningar av SBR-tekniken. Vidare beskrivs nya tekniker, metoder och förbättringar i de tillämpningar där SBR-teknikens utnyttjas. Uppfinningen kan användas inom valfria ljudkodare samt till olika former av talkodare och vid konstruktion av nya kodare. Tekniken implementeras i hård- eller mjukvara. Förbättring av kodningstekniken innebär reducerat bitflöde och förbättrad ljudkvalitet.

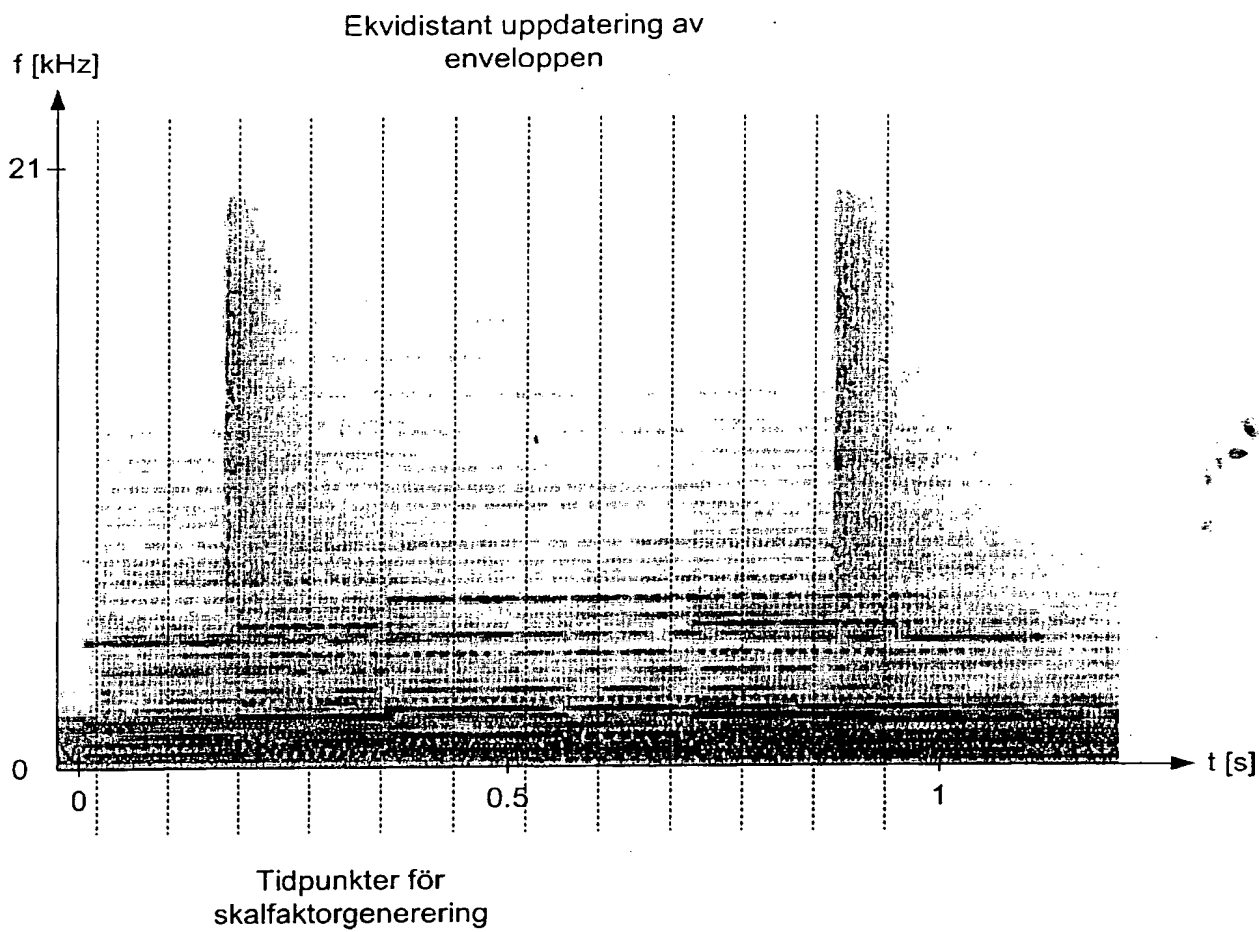


FIG 1

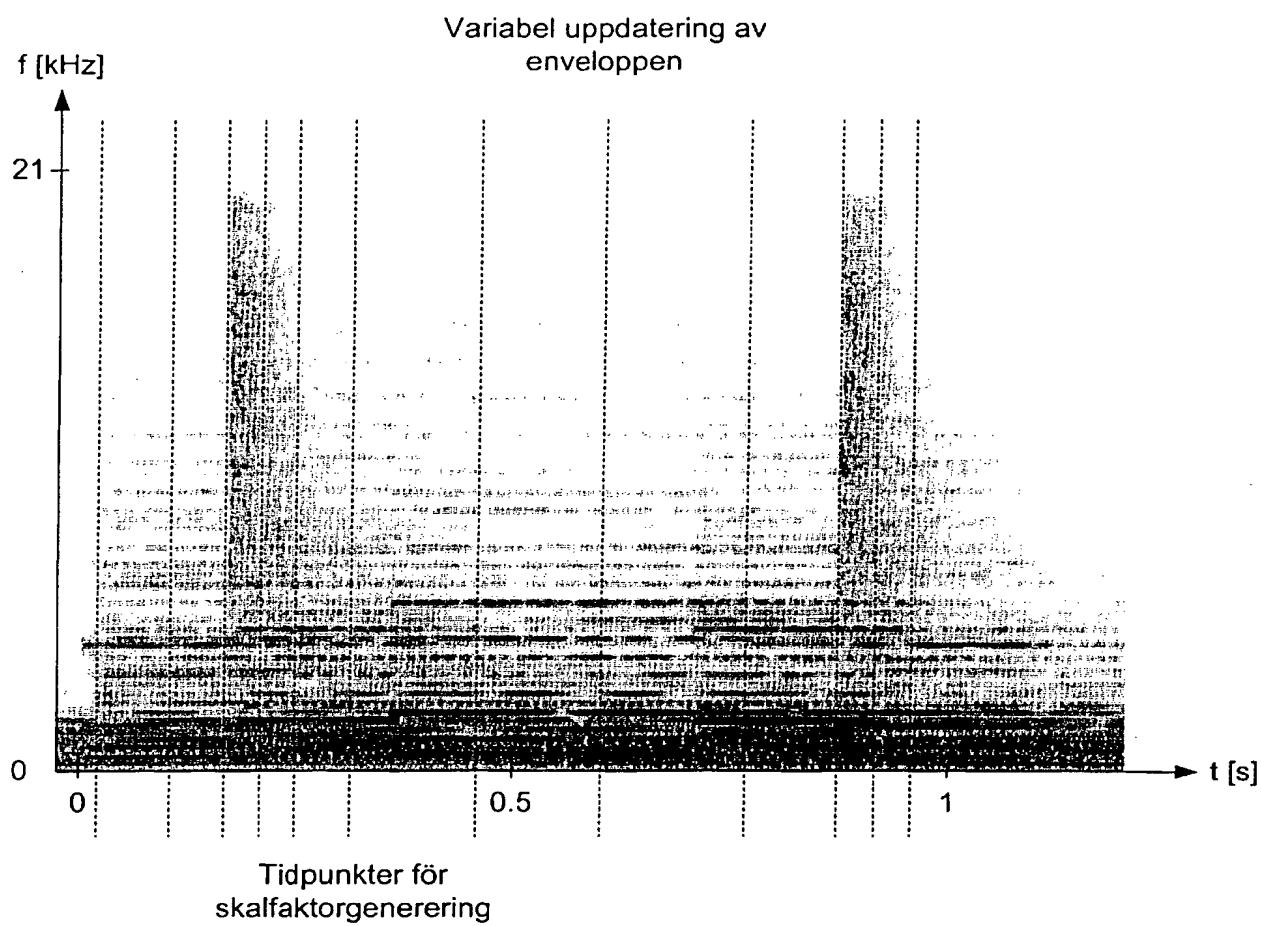
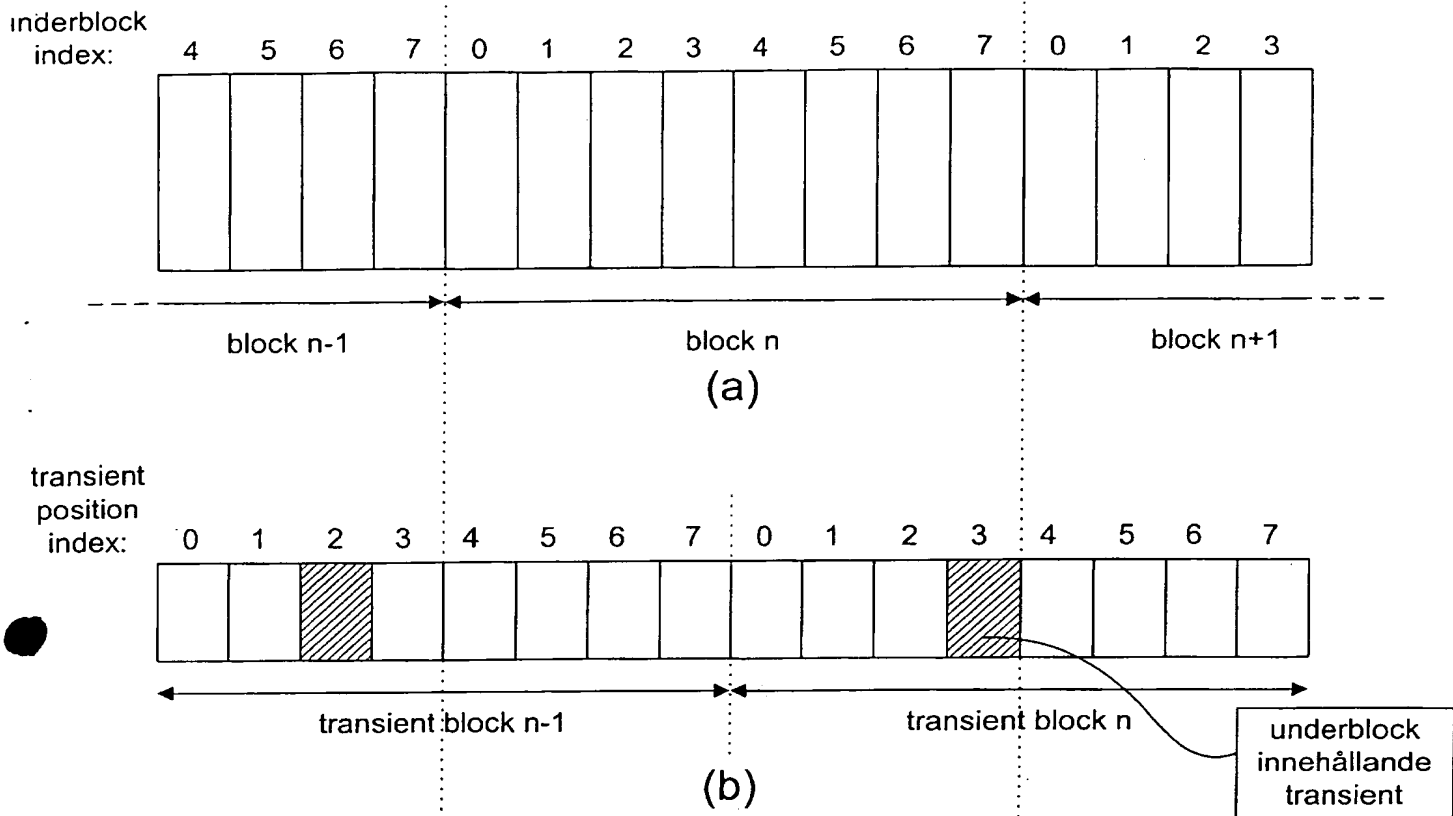


FIG 2

3



Föregående transient position: 2,  
nuvarande transient position: 6  
ger följande tid/frekvens-indelning  
av block n ur tabell:



Låg  
frekvensupplösning

Hög  
frekvensupplösning

Låg  
frekvensupplösning

(c)

FIG 3

4

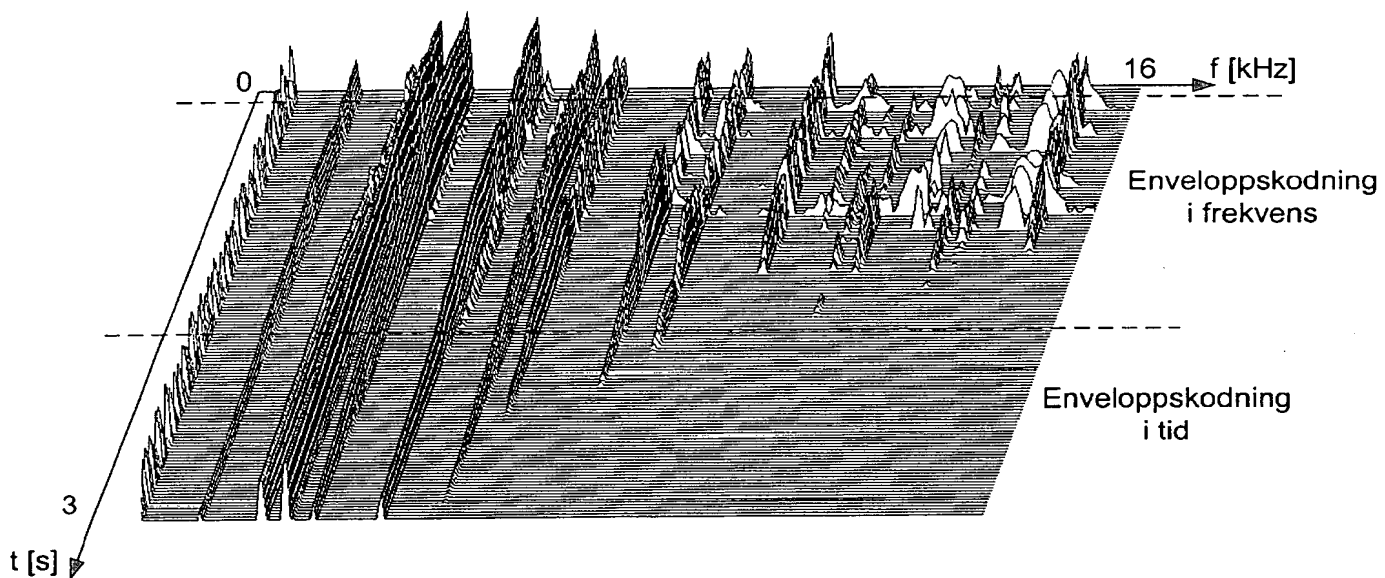


FIG 4



5

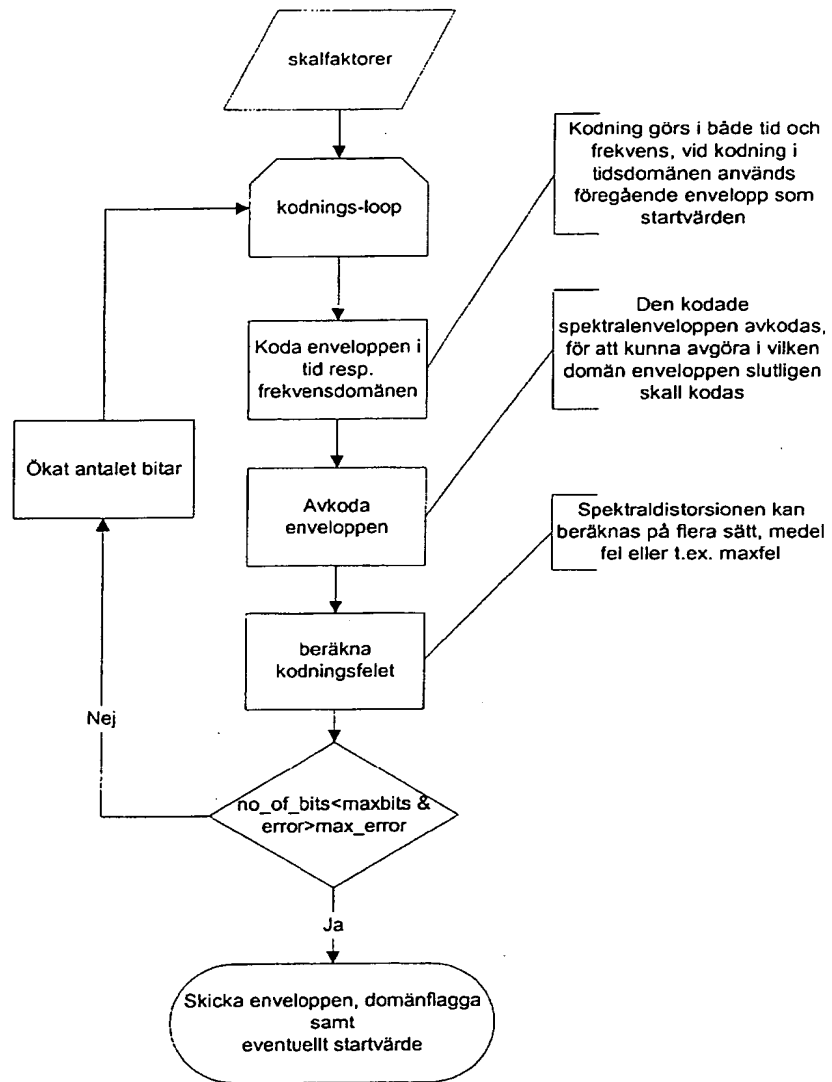
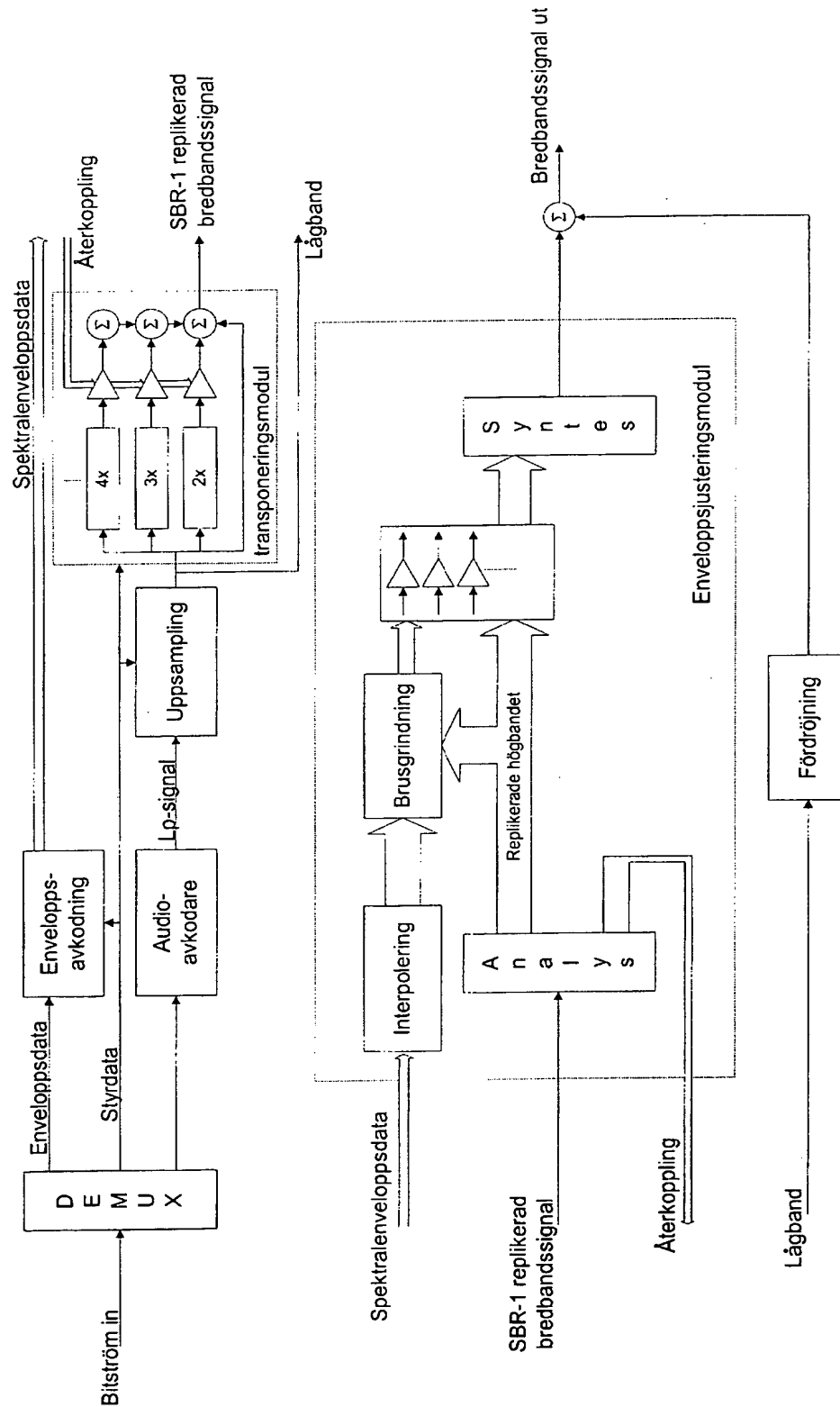


FIG 5

2010-06-14

FIG 6



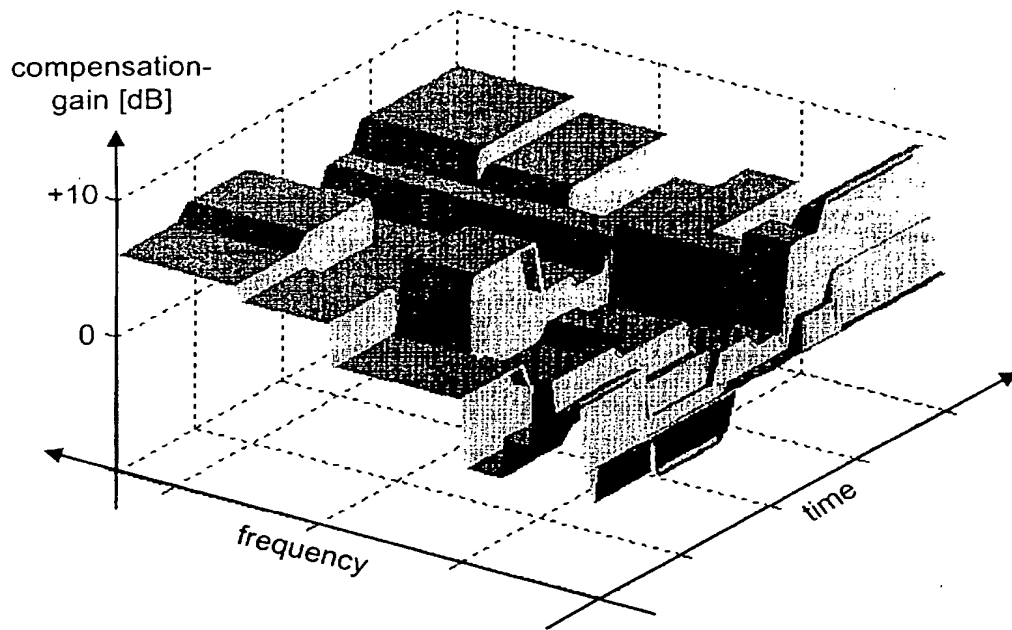


FIG 7A

8

rn030127

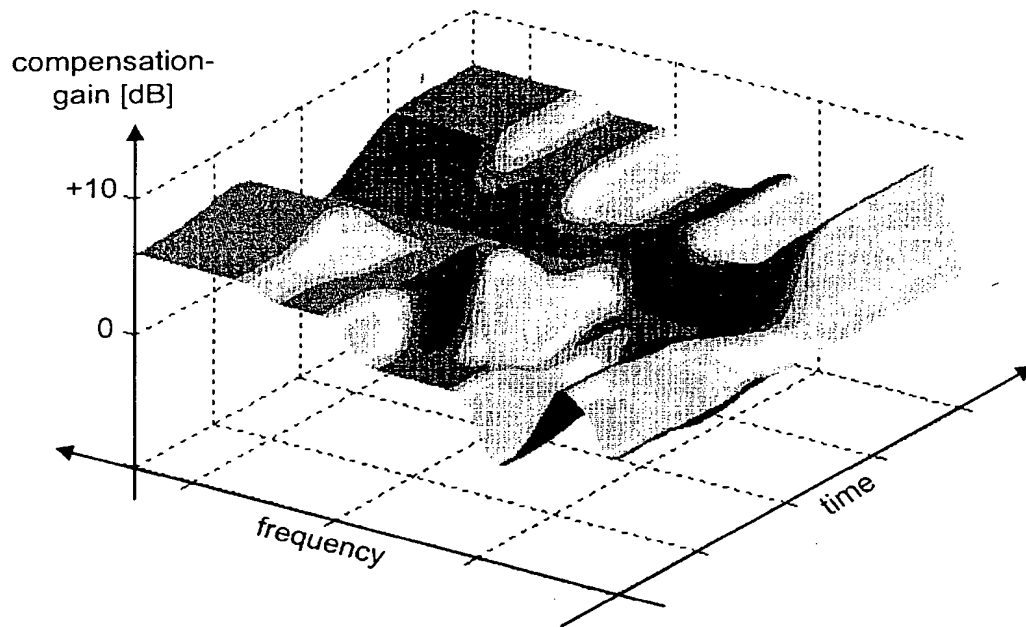


FIG 7 B

9

10107

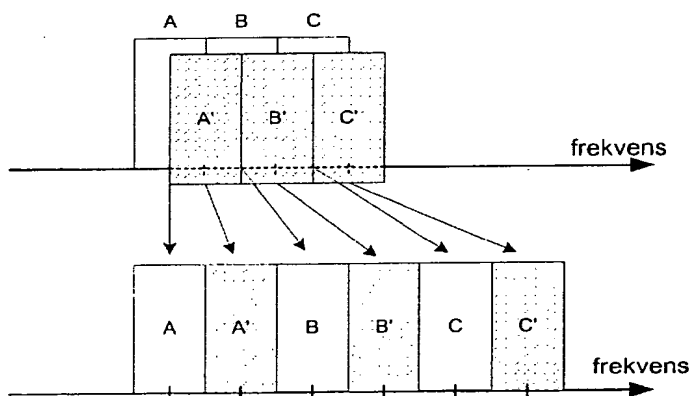
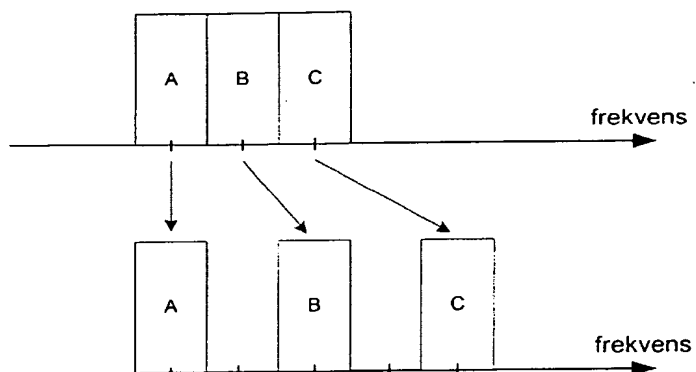


FIG 8

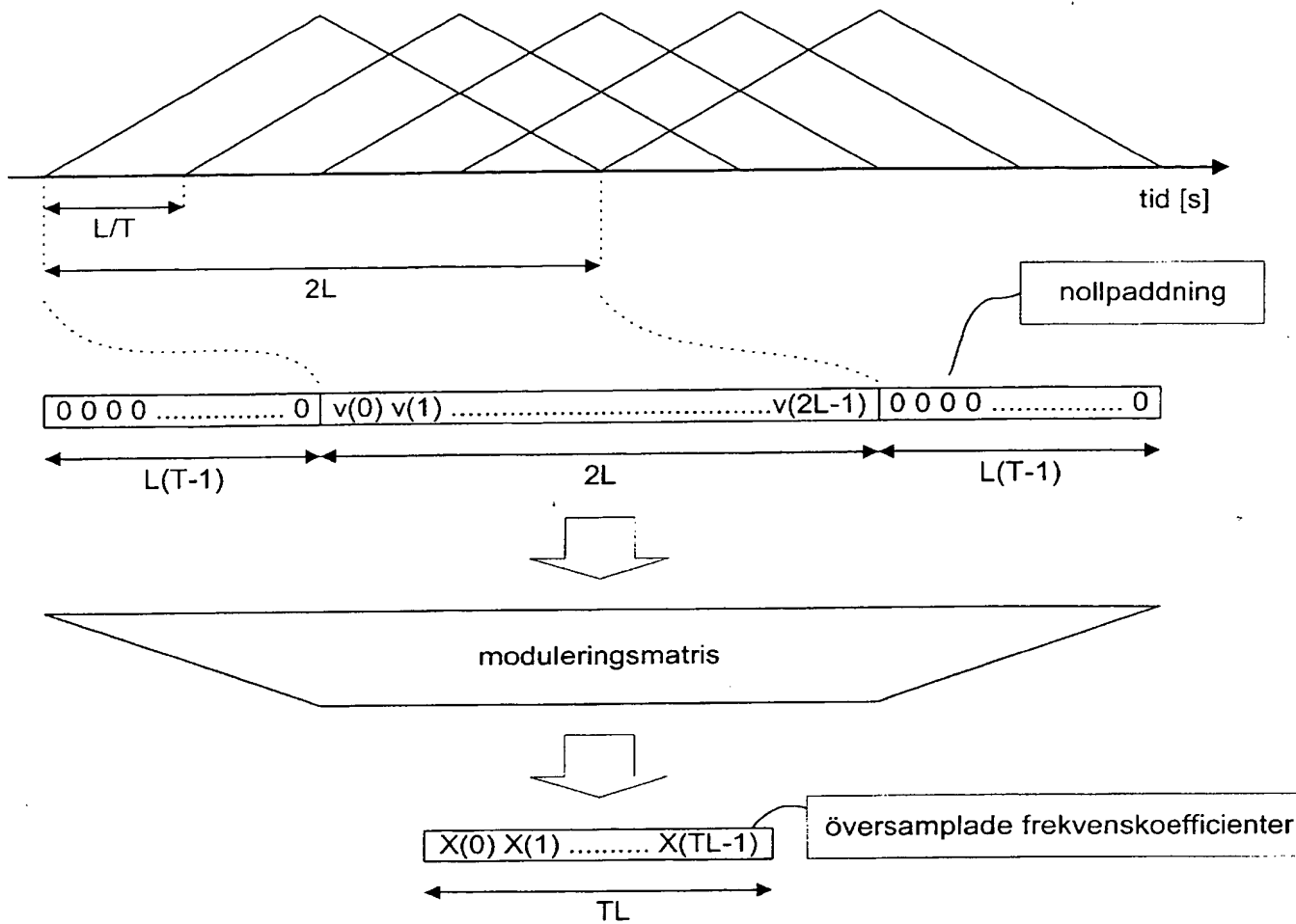


FIG 9

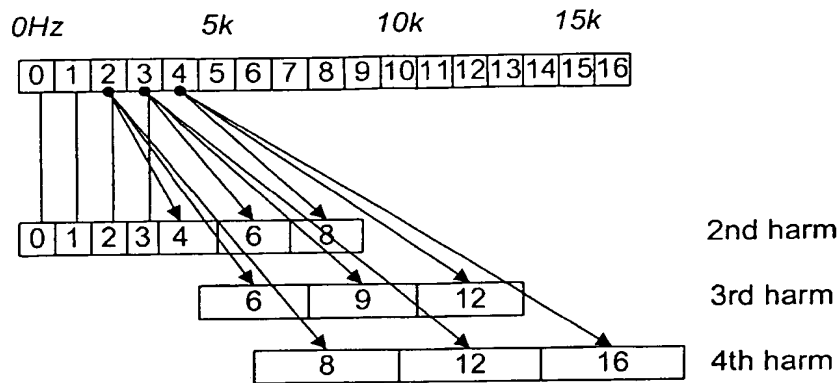


FIG 10A

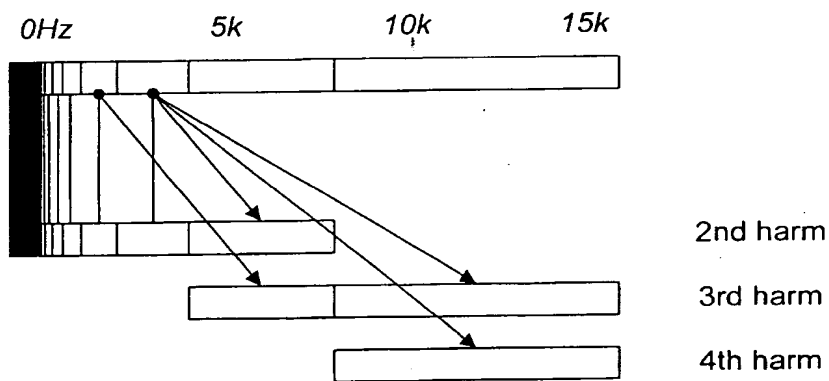


FIG 10B

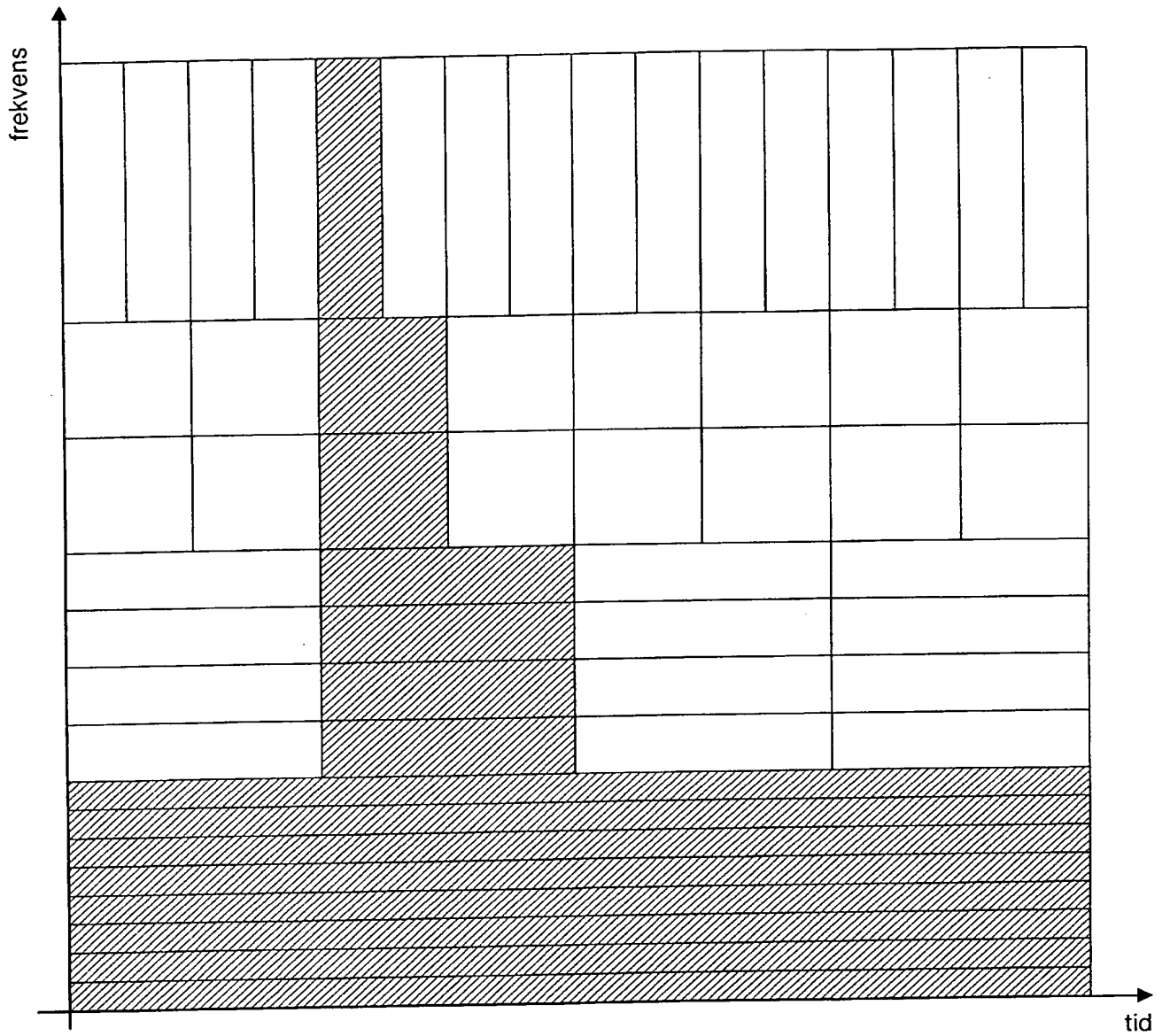


FIG 11



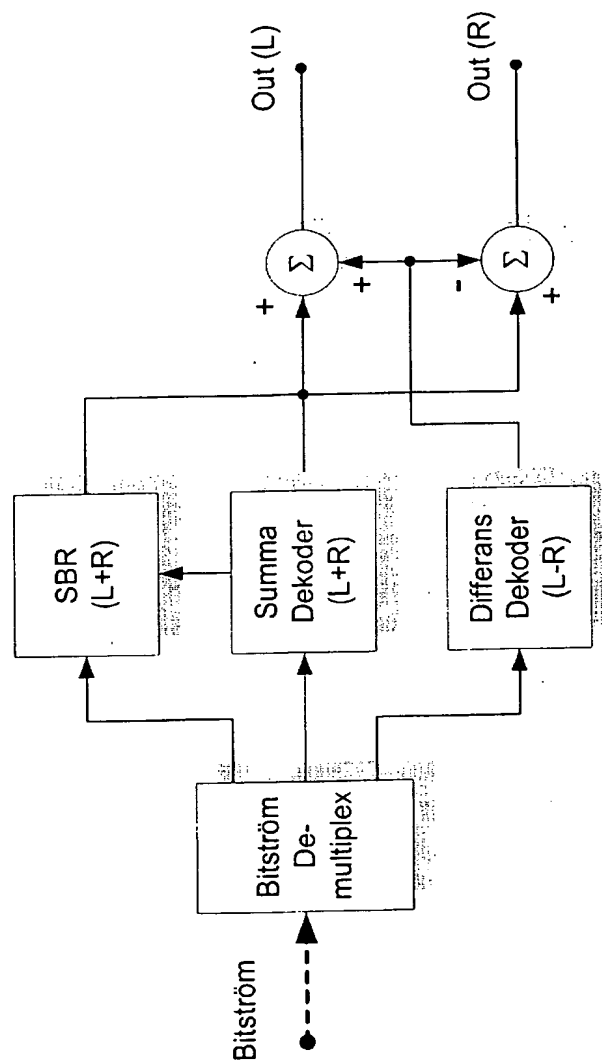


FIG 12

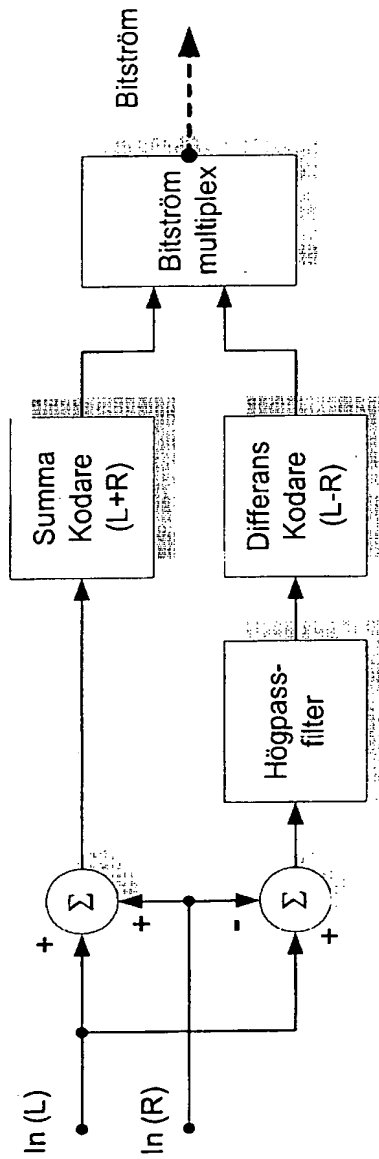


FIG 13

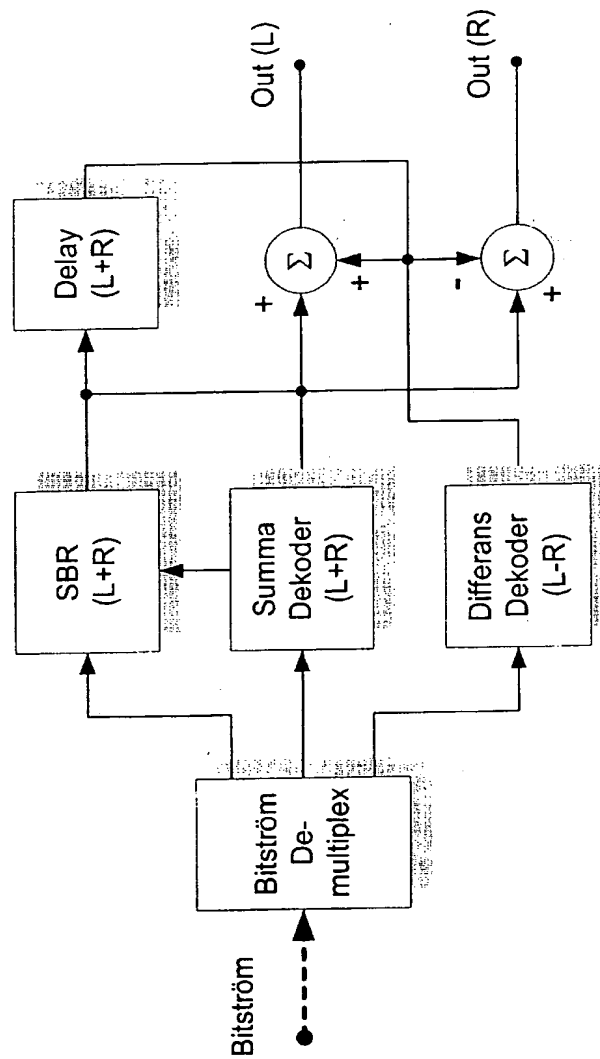


FIG 14

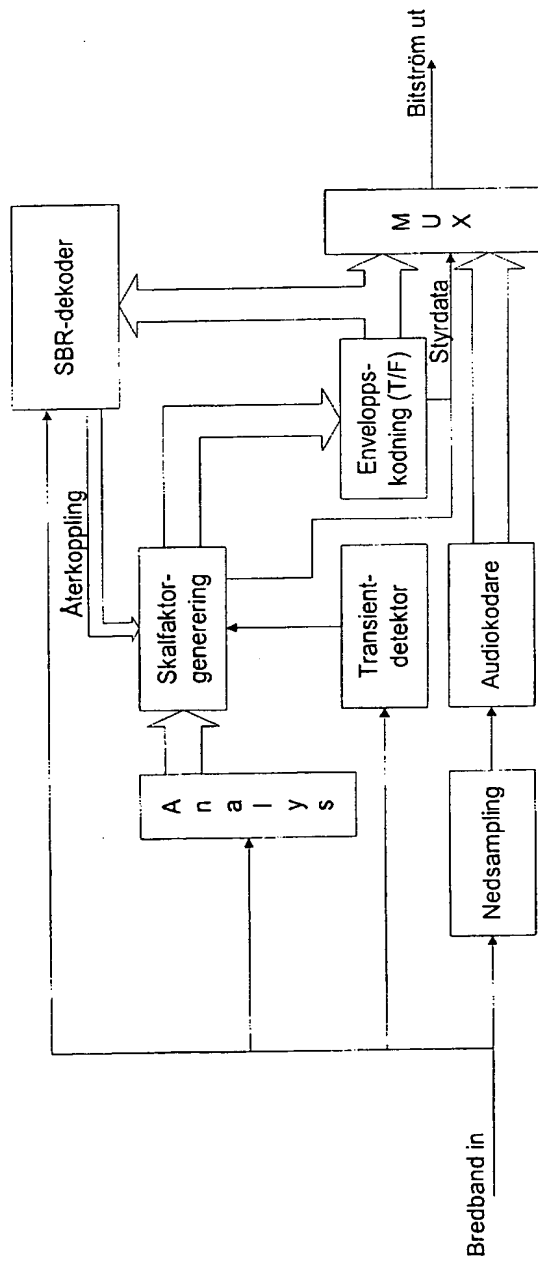


FIG 15